Manuscript
Journal of Experimental Psychology: Learning, Memory, & Cognition, in press

Updating Self-Location by Self-motion and Visual Cues in Familiar Multiscale Spaces

Xuehui Lei, Weimin Mou

Department of Psychology, University of Alberta

Author's Note

Correspondence concerning this article should be addressed to Xuehui Lei or Weimin Mou, P217 Biological Sciences Bldg., Department of Psychology, University of Alberta, Edmonton, Alberta, Canada, T6G 2E9. Contact: xuehui1@ualberta.ca or wmou@ualberta.ca

Abstract

This study examined functions of self-motion and visual cues in updating people's actual headings in multiscale spaces. In an immersive virtual environment, the participants learned objects' locations inside two misaligned rectangular rooms by locomoting within and between the rooms. In each testing trial, the participants locomoted to adopt an actual perspective in one room, and then they judged relative direction to a target from an imagined perspective in the other room (remote perspective-taking). The imagined and actual perspectives had the same/opposite cardinal directions (globally aligned/misaligned) or had the same/opposite orientations defined by room structures (locally aligned/misaligned). Global/local sensorimotor alignment effects mean that performance is better when imagined and actual perspectives were globally/locally aligned than misaligned. We examined these effects to infer updating actual headings in global/local representations. The results showed local but no global sensorimotor alignment effect. By contrast, there were both global and local sensorimotor alignment effects when the participants judged across-room relative headings prior to remote perspective-taking. These results indicate that people update headings in local representations based on visual similarities between local spaces. People update headings in global representations based on self-motion cues available in across-boundary navigation, but updating headings globally requires tasks to activate global-relevant sensorimotor representations.

Keywords: sensorimotor alignment effect; multiscale spatial representations; path integration; piloting; spatial updating

Updating Self-Location by Self-motion and Visual Cues in Familiar Multiscale Spaces

1. Introduction

In navigation, it is essential for people to update their self-location in the environment, including their locations (positions) and orientations (headings). Two methods can be used in self-localization. One method is path integration, using self-motion cues during locomotion to continuously update self-location (Etienne & Jeffery, 2004; Loomis, et al., 1999; Mittelstaedt & Mittelstaedt, 1980; Wang, 2017; Yamamoto & Shelton, 2005). The other method is piloting, using visual cues (e.g., familiar landmarks) to intermittently update self-location (Cheng & Spetch, 1998; Etienne et al., 2004; Foo et al., 2005; Wehner et al., 1996). Since these two methods rely on bottom-up information (self-motion and visual cues), we refer to them as *bottom-up updating* of *actual* self-location. Other than bottom-up updating, people can also take imagined perspectives in a familiar space without using bottom-up information (e.g., being blindfolded and disoriented). As this mental perspective-taking creates imagined self-location only using top-down information (i.e., spatial memory of the familiar space), we refer to it as *top-down updating* of imagined self-location.

There are empirical evidence indicating bottom-up updating of self-location in immediate environments. After disorientation in a familiar environment, visual cues including extended surfaces (e.g., boundaries) and landmarks (e.g., a traffic cone) can update self-location (Cheng, 1986; Doeller & Burgess, 2008). Path integration also updates self-location (Attneave & Farrar, 1977; Klatzky et al., 1998; May & Klatzky, 2000; Rieser et al., 1986; Waller et al., 2002). While people have difficulty adopting an imagined perspective different from their actual perspective, they have no difficulty to adopt this imagined perspective after physically rotating to be aligned with it without vision (Rieser, 1989).

People also navigate multiscale environments (Brockmole & Wang, 2002; Han & Becker, 2014). A multiscale environment (an environmental space, Montello, 1993) contains several individual spaces (a vista space, Montello, 1993) that are separated by boundaries (e.g., walls). Studies have examined multiscale spatial representations consisting of local representations of spatial relations within individual spaces and global representations of spatial relations between spaces (Brockmole & Wang, 2002; Han & Becker, 2014; Lei et al., 2020; Marchette et al., 2014; Shine et al., 2016). The findings indicate limitations of developing multiscale spatial representations. Marchette et al. (2014) showed that after the participants learned a virtual environment with misaligned museums by navigating within and across museums, they only developed local representations of individual museums but no global representations of spatial relations between two objects across museums. Lei et al. (2020) further showed that local and global representations coexisted only after the participants had learned the global environment prior to learning objects in individual spaces.

However, it is not clear to what extent bottom-up updating of self-location occurs in familiar multiscale spaces with the support of multiscale spatial representations. One may speculate that path integration and piloting update self-location in multiscale spaces just as in immediate spaces (e.g., Doeller & Burgess, 2008; Rieser, 1989). Yet, previous studies have indicated that people can automatically update their headings using path integration only in a space with sensorimotor interaction (e.g., a directly viewed or touched space), but not in a space without sensorimotor interaction (e.g., an imagined space, a verbally described space) (Avraamides, 2003; Wang, 2004; Wang & Brockmole, 2003a; but see Loomis, et al., 2007). In multiscale spaces, people do not seem to have sensorimotor interaction with all spaces due to the blockage by boundaries. Consequently, we should be cautious to generalize the findings of updating self-location by piloting and path integration in immediate spaces to multiscale spaces.

To foster the theoretical understanding of human spatial memory and navigation, the current study examined bottom-up updating during navigation in a familiar multiscale space with the support of multiscale spatial memories. Specifically, we investigated how people use visual (piloting) and self-motion (path integration) cues to update their headings relative to a remote room, while standing in the immediate room with minimal visual cues outside[1].

Based on visual similarity between the immediate and remote room, piloting may update people's headings relative to the remote room even with minimal visual cues outside the immediate room. Riecke and McNamara (2017) showed that after the participants learned objects' locations in a learning room and were disoriented to be led to a testing room, they re-anchored self-location in the learning room based on similar room structures of the learning and testing rooms. Thus, piloting updates self-location relative to a remote room via visual-based re-anchoring. This visual-based re-anchoring is similar to the daily-life experience that people can know their self-location in a Walmart store of a new city, because they have visited Walmart stores in their home city and different Walmart stores presumably have identical spatial layouts. As the visual-based re-anchoring is determined by similarities of room structures, global spatial relations between the two room are not required when piloting updates self-location relative to the remote room. We refer to this means of updating as *visual-re-anchoring.*

Path integration may also update people's headings relative to the remote room. For example, when we move from the bedroom to the bathroom at home in darkness, we in the

[1] It is obvious that people can use piloting to update their global headings if they can directly see global landmarks. What is less obvious is how people update headings relative to a remote space when boundaries block views outside the immediate space. Thus, we minimized the global landmarks in the current study.

bathroom may still know our headings relative to the bedroom. In contrast to *visual-re-anchoring*, path integration updates self-location relative to the remote room relying on the global representations of the spatial relations between the immediate and remote rooms. We refer to this means of updating as *global-path-integration*.

Therefore, updating headings relative to the remote room using visual and self-motion cues could occur by means of *visual-re-anchoring* and *global-path-integration*. These two means can be demonstrated by two different sensorimotor alignment effects when participants take actual perspectives in an immediate space (with both visual and self-motion cues) and adopt imagined perspectives in a remote space (remote perspective-taking, Riecke & McNamara, 2017). A sensorimotor alignment effect is defined by better performances in mental perspective-taking when imagined headings and actual headings are the same than when these two headings are different (Avraamides & Kelly, 2008; Kelly et al., 2007; Mou et al., 2004). It is attributed to the congruency between imagined heading representations from top-down updating (relying on top-down information only) and actual heading representations from bottom-up updating. When imagined headings and actual headings are different, these two heading representations are incongruent, thus interfering with the mental perspective-taking. When imagined headings and actual headings are the same, these two heading representations are congruent, thus not interfering with the mental perspective-taking.

Specifically, in remote perspective-taking, one sensorimotor alignment effect is better performances when the imagined and actual perspectives are aligned according to local visual similarity (e.g., both facing the doors). This local sensorimotor alignment effect demonstrates that participants re-anchor self-location in the remote space using visual similarity (*visual-re-anchoring)*. The other one is better performances in remote perspective-taking when the imagined and actual perspectives are aligned according to global spatial relations (e.g., facing the same cardinal direction). This global sensorimotor alignment effect indicates that participants update self-location using *global-path-integration.*

The current study proposed three hypotheses stipulating whether and how people use either of these two means: the *visual-re-anchoring only* hypothesis, the *dual means* hypothesis, and the *conditional dual means* hypothesis.

The *visual-re-anchoring only* hypothesis claims that bottom-up updating relative to a remote space only occurs by means of *visual-re-anchoring* but not by means of *global-path-integration.* According to this hypothesis, local but no global sensorimotor alignment effect occurs in remote perspective-taking. Previous studies have shown local sensorimotor alignment effects in remote perspective-taking (e.g., Riecke & McNamara, 2017). However, to the best of our knowledge, no study has shown a global sensorimotor alignment effect (Kelly et al., 2007; Liu & Xiao, 2018; May, 2007; Riecke & McNamara, 2017; Shelton & Marchette, 2010). It is worth noting that in most of the previous studies (Liu & Xiao, 2018; May, 2007; Riecke & McNamara, 2017; Shelton & Marchette, 2010), participants were disoriented on purpose between spaces removing the possibility of *global-path-integration*. In an exceptional case, the participants in Kelly et al. (2007) walked a few meters on a straight path between the learning and testing rooms. They seemed to have some global spatial memories as they could point accurately to the learning room from the testing position (in their Experiment 1), but the results still showed no global sensorimotor alignment effect. The null global sensorimotor alignment effect, especially when people seem to have developed global spatial memories, suggests that path integration may not be able to access or use global spatial memories. When people switch between spaces across boundaries, they may lose track of the previous space and only track the immediate space (Wang, 2016; Wang & Brockmole, 2003b).

However, other theoretical proposals and empirical evidence suggest that path integration interacts with global spatial memories. Theorists propose that path integration is critical to develop global spatial memories when piloting cues are rare in large-scale spaces (Gallistel, 1990; Jacobs & Schenk, 2003; Loomis et al., 1999; McNaughton et al., 2006; Meilinger, 2008). Previous studies have also indicated that people can develop global spatial memories by across-boundary navigation (Lei et al., 2020; Shine et al., 2016). As piloting is unlikely to function globally across boundaries due to boundaries blocking views, path integration should be the primary method to integrate representations of across-boundary spaces (Jacobs & Schenk, 2003). Moreover, as path integration is not impaired by crossing boundaries (Mou & Wang, 2015), path integration is able to access global spatial memories of across-boundary spaces.

Thus, we propose the *dual means* hypothesis stipulating that both means of *global-path-integration* and *visual-re-anchoring* occur in bottom-up updating relative to a remote space. According to this hypothesis, people show both global and local sensorimotor alignment effects in remote perspective-taking. This hypothesis attributes the null global sensorimotor alignment effect in the previous studies to the fact that participants did not develop global spatial memories of across-boundary spaces (Riecke & McNamara, 2017; Shelton & Marchette, 2010). This hypothesis treats Kelly et al. (2007) as a single exceptional case.

The *conditional dual means* hypothesis is a variant of the *dual means* hypothesis. It states that people can update headings relative to a remote space by both means, but *global-path-integration* requires sensorimotor representations of the global environment (sensorimotor global representations). People perceive and locomote in the immediate space without sensorimotor engagement with a remote space (e.g., we cannot see or collide with objects in another room). Hence, sensorimotor spatial representations in working memory are primarily concerned with the immediate local space (Wang, 2004; Wang & Brockmole, 2003a), though there are both global and local spatial representations in long-term memory. Without sensorimotor global representations, path integration cannot globally update self-location and the means of *global-path-integration* will not be used.

Nonetheless, global spatial representations might become accessible on the sensorimotor level in some situations. Sholl et al. (2006) (see also Burte & Hegarty, 2014) showed that knowing ones' headings on campus by looking out the window could affect performances in judging the allocentric headings indicated by campus photographs, suggesting that global spatial representations can be activated on the sensorimotor level. Previous studies also showed that the types of updated spatial representations varied with instructions (He & McNamara, 2018; Wiener et al., 2011). Inspired by these findings, we speculate that although local representations are the primary type of sensorimotor representations, global representations can also be accessed on the sensorimotor level when global representations are activated by visual cues of the global environment or emphasized by task requirements (e.g., explicitly requiring participants standing in the immediate room to face the global direction of a probed heading in the remote room). With sensorimotor global representations, the means of *global-path-integration* will be used.

Hence, a local sensorimotor alignment effect occurs in remote perspective-taking whereas a global sensorimotor alignment effect does not occur unless global spatial representations are activated and become sensorimotor global representations. The null global sensorimotor alignment effect in the previous studies might be due to (a) no sensorimotor global representations (Kelly et al., 2007) or (b) no global spatial representations in long-term memory (Riecke & McNamara, 2017; Shelton & Marchette, 2010).

Overall, all three hypotheses predict a local sensorimotor alignment effect due to the means of *visual-re-anchoring* (Riecke & McNamara, 2017), but they differ in predicting a global sensorimotor alignment effect caused by the means of *global-path-integration*. Three experiments were conducted to differentiate these hypotheses by examining global sensorimotor alignment effect as well as the local sensorimotor alignment effect.

2. Experiment 1

Experiment 1 examined the global and local sensorimotor alignment effects in remote perspective-taking, when both multiscale spatial representations had been formed. It is critical to ensure multiscale representations, especially global representations, in long-term memory, since null global representations would lead to a null global sensorimotor alignment effect (Riecke & McNamara, 2017; Shelton & Marchette, 2010). Thus, we adopted the environmental setup and the learning procedure in Lei et al. (2020, Experiment 3)[2], which showed evidence of multiscale spatial representations after learning an immersive virtual environment containing two rooms with similar local structures but misaligned global orientations. However, the testing phase differed from that in Lei et al. (2020). Lei et al. (2020) tested participants' remote perspective-taking in a physical room outside the virtual rooms so there were no valid self-motion or visual cues associated with the virtual environment. Contrarily, in the remote perspective task of the current study, on each testing trial, the participants took an actual heading in one of the learned virtual rooms with both self-motion and visual cues available and conducted mental perspective-

[2] To ensure that the learning phase adopted from Lei et al. (2020) could lead to global and local representations, we conducted another experiment with the exact same experimental setup and the learning procedure in the virtual environment and the mental perspective-taking task in a structurally different experimental room. The results replicated the findings of global and local priming effects (see supplementary materials).

taking in the other virtual room. Hence, the current study is concerned with bottom-up updating by testing the congruency between the self-location representations from bottom-up updating and top-down updating whereas Lei et al. (2020) only examined the top-down updating.

Specifically, on each testing trial, the participants moved to face an actual view in one room (an immediate room) but imagined facing a view in the other room (a remote room), and then pointed to a target in the remote room from the imagined perspective. We manipulated the actual and imagined views to be globally/locally aligned/misaligned. All three hypotheses predict a local sensorimotor alignment effect. While the *visual-re-anchoring only* hypothesis predicts a null global sensorimotor alignment effect, the *dual means* hypothesis predicts a global sensorimotor alignment effect. The *conditional dual means* hypothesis leans toward the prediction of no global sensorimotor alignment effect since this experiment did not manipulatively activate global spatial representations.

2.1 Method

2.1.1 Participants

The study was approved by the Ethics Committee of the University of Alberta. Forty-eight university students (24 females) with normal or corrected-to-normal vision participated to partially fulfill the requirement for an introductory psychology course.

In Experiment 3 of Kelly et al. (2007), Cohen's *d* of the local sensorimotor alignment effect in remote perspective-taking was about 1.07[3], and partial eta squared was about 0.38[4]. Because the remote perspective-taking task in the current study had fewer trials than that in

---

[3] Cohen's *d* was calculated by $\sqrt{(\frac{2F}{N})}$. In Experiment 3 of Kelly et al. (2007), the *F* value for the sensorimotor alignment effect when testing in a novel environment was 9.19 and *N* was 16.

[4] $\eta^2_p = \frac{dfe \times F}{dfi + dfe \times F} = \frac{Nd^2}{2(N-1)+Nd^2}$, where *dfe* and *dfi* are the *df* of error (denominator) and within (nominator) in the *F* test respectively. Particularly, *dfe*=1, *dfi*=(*N*-1), *N* is the number of observers in each group, *d* is Cohen's *d* . When *N* is large, $\eta^2_p = \frac{dfe \times F}{dfi + dfe \times F} = \frac{d^2}{2+d^2}$.

Kelly et al. (for each of two blocks, there were 12 trials in the current study versus 56 trials in Kelly et al.), the effect size should decrease significantly. Assuming partial eta squared would decrease to 0.113, which corresponded to the effect size decreasing to 0.50 (a medium effect), we used 48 participants to get the power value of 0.67 (for within-subject design with a two-level variable, see the Matlab code for the power analysis at https://doi.org/10.7939/r3-aqm4-3p16).

2.1.2 Materials and design

The experimental lab space was a square room that was 4 m by 4 m. The immersive virtual environment was presented using Vizard software (WorldViz, Santa Barbara, CA) in a head-mounted display (HMD, Oculus Rift, Oculus VR, LLC., Irvine, CA). The participants' head motions were tracked by an InterSense IS-900 motion tracking system (InterSense, Inc., Massachusetts), so that they could physically turn to change viewing orientations in the virtual environment. The participants used a moving stick on a gamepad to move along their viewing orientation in the virtual environment. During learning when the participants were asked to point to a direction, they used as a pointer a virtual blue stick associated with an InterSense Wand. In the remote perspective-taking task during testing, the participants used a joystick (Logitech Extreme 3D Pro, Newark, CA) to judge the relative direction to a target from an imagined perspective.

The center of the virtual environment overlapped with the center of the experimental lab room. The virtual environment had a grassy ground with distal orientation cues (i.e., ocean, mountain, forest, and city) in the open field (Figure 1). There were two rectangular rooms with a 90° angular difference. Before seeing the two rooms, the participants learned five other buildings. The buildings and the two rooms were never presented simultaneously. When learning the five buildings, the participants only saw the doorways of the two rooms. When the participants started to learn objects in the two rooms, the five buildings were not presented. Learning the five buildings with distal cues was to support the development of global representations of the two rooms (Kelly & McNamara, 2010; Lei et al., 2020; Philbeck & O'Leary, 2005).

The two rooms had identical geometry with an aspect ratio of 0.6 (36m by 60m). Both rooms had a door, a window on the back wall, and a long carpet from the door to the back wall. Each room had distinguishable interior and exterior colours and textures. In every corner of the rooms, there was a stage on the ground and two alcoves by the walls with an object in each alcove. The alcove had a door so that the participants could not see the object in it unless they stood on the corresponding stage in the corner to make the door disappear. Thus, for each view (i.e., viewing an object), the participants had a fixed standing position (i.e., the corresponding stage) and facing direction (i.e., facing the alcove while standing on the stage). Specifying each view would determine a standing position and a facing direction.

[Figure 1]

For all trials of the remote perspective-taking task, participants had the actual view in the immediate room and the imagined view in the remote room. The participants pointed to a target in the remote (imagined) room from the imagined view (e.g., "Imagine you are facing the mug," "point to the kettle"). The imagined views were the initial views of the two rooms (i.e., Views 1, 2, 9, and 10), which were the first views that the participants saw when entering the rooms. We focused on the initial views because previous studies have shown that priming effects of the same view in perspective-taking (compared with a different view) only occurred for the initial views (Avraamides & Kelly, 2005; Lei et al., 2020).

We manipulated the relationships between the actual views and the imagined views. There were four types of trials: globally aligned, globally misaligned, locally aligned, and locally misaligned. The first two were contrasted to test the global sensorimotor alignment effect whereas the last two were contrasted to test the local sensorimotor alignment effect. Table 1 shows all pairs of the actual and imagined views for the four trial types when the participants had actual views in Room 2 shown in Figure 1.

[Table 1]

In the globally aligned trials, the angular difference between the actual and imagined views was globally 0° (e.g., actual view being View 15 and imagined view being View 2 were both facing the ocean); whereas in the globally misaligned trials, the angular difference was globally 180° (e.g., actual view being View 12 was facing the city but imagined view being View 2 was facing the ocean). In both globally aligned and globally misaligned trials, the local angular difference between the actual and imagined views was 90° (e.g., actual view being View 12 was facing a side wall in one room but imagined view being View 2 was facing the window in the other room), removing any confound from the local sensorimotor alignment effect. In the locally aligned trials, the angular difference between the actual and imagined views was locally 0° (e.g., actual view being View 10 and imagined view being View 1 were both facing the windows in the two rooms); whereas in the locally misaligned trials, the angular difference was locally 180° (e.g., actual view being View 13 was facing the door in one room but imagined view being View 1 was facing the window in the other room). In both locally aligned and locally misaligned trials, the global angular difference between the actual and imagined views was 90° (e.g., actual view being View 13 was facing the forest but imagined view being View 1 was facing the ocean), removing any confound from the global sensorimotor alignment effect.

For each type of trial, three trials were generated by choosing three different targets in the imagined room. The target objects were randomly chosen from all objects in the imagined room except for the viewing object and the object on their exact back (e.g., when the participants imagined facing View 1, View 6 was on their exact back). Consequently, for each imagined room, there were 12 trials. Both rooms were used as the imagined room alternatively and trials using the same imagined room were administered within a block (producing 24 trials in total, six for each type of trial, see sample trials at https://doi.org/10.7939/r3-ch2r-xr18). The order of the two blocks was counterbalanced across the participants and the order of the 12 trials within each block was randomized.

2.1.3 Procedure

Before the experiment, the participants signed consent forms, read instructions and practiced how to use the joystick to point. Then the participants were blindfolded and guided to the experimental lab room. They stood at the center of the experimental room, which was also the center of the virtual environment. They were required to close their eyes, remove the blindfold and put on the HMD.

2.1.3.1 Learning buildings

In the learning phase, the participants first stood at the center of the virtual environment and saw an open field of grass-textured ground with distal orientation cues. Then the doorways of the two rooms (without the rooms) were presented. When the participants went to stand in the middle of one doorway, the five buildings were presented. The participants learned the buildings' directions for one minute and the buildings disappeared. The participants used a pointer to indicate the original direction of the buildings. After each pointing response, the probed building was presented at the original place as feedback. There were two blocks to indicate the buildings' directions. In each block, the buildings were tested randomly. After learning the buildings from one doorway, the buildings were presented. The participants moved to the middle of the other doorway to learn and then be tested about the buildings' directions. The order of the two doorways was counterbalanced across the participants. After the participants learned the buildings' directions from the two doorways, a pole appeared at the center of the virtual environment. The participants moved to the pole and the pole disappeared when the participants met the pole. Then the buildings disappeared and the two rooms were presented.

Learning the five buildings with distal cues, which was the same procedure used in Lei et al. (2020), was to encourage developing global representations of local rooms. Some prior spatial learning of the global environment might provide a global framework to scaffold spatial information in local rooms and integrate them into larger/newer global representations (Kelly & McNamara, 2010; Philbeck & O'Leary, 2005). The buildings would not be presented or tested in the testing phase.

2.1.3.2 Learning locations of objects within rooms

The participants learned the objects' locations in both rooms. The learning order of the two rooms was the same as the order of the two doorways when learning the buildings. After the participants entered one room, the experimenter introduced the names of the objects by instructing the participants to move from the left corner under the window of the room and then to move clockwise to the next corner. To view an object, the participants moved onto the corresponding stage and faced the alcove to see the object in it. After familiarizing themselves with the objects' names, the participants moved freely in the room to learn the locations of the objects. Five minutes later, the objects disappeared and the participants replaced the objects to the original alcoves. A probed object with its name was presented in the center of HMD. To replace an object, the participants moved onto the corresponding stage and faced the correct alcove, and then they pressed a button to place the object back. The object appeared in the correct alcove as feedback. If the participants were wrong, the experimenter would direct them to the correct location to see the object.

There were two blocks to replace the objects and the objects were tested randomly in each block. After learning in one room, the participants moved to the other room and repeated the same learning procedure as in the first room. Finally, the participants were asked to replace all the objects in the two rooms. Each object was tested once with feedback. Since the objects were tested randomly, the participants had to move between the rooms to replace them. The experimenter recorded the number of errors when the participants replaced the object into a wrong alcove[5]. After the learning phase, the participants were given a five-minute break.

2.1.3.3 Testing of remote perspective-taking

The participants entered the first testing room and then the experimenter pressed a key to start the block of trials. When the block in one room started, the other room disappeared to ensure that the participants could not visually perceive the other room.

In each trial, a sentence was presented at the center of the screen on the HMD to instruct the participants to move and face an actual view (e.g., “Move to face the hat”). After the participants moved onto the corresponding stage, the correct object was presented. If they moved onto a wrong stage, the participants were directed to move to the correct stage. Facing the correct object, the participants pressed a button on the joystick to continue this trial. The participants

[5] The number of errors in all experiments were quite low. Means of percentage errors were 3.39%, 2.47%, and 2.93%. Standard deviations were 5.15%, 5.27%, and 4.49%. These results suggest that the participants learned well about the objects’ locations in the two rooms.

were required to keep their actual facing direction and were not allowed to turn around. As soon as they pressed the button, a sentence with an imagined view in the other room was presented (e.g., "Imagine you are facing the mug"). The participants were told to keep their actual facing direction while mentally taking the imagined perspective in the remote room. When the participants took the imagined perspective, they clicked a trigger on the joystick. The duration between the appearance of the imagined view and clicking the trigger was termed orientation latency. After the participants clicked the trigger, a sentence was presented with instructions to point to a target object (e.g., "point to the kettle"). The participants used a joystick to point. They were required to respond as quickly as possible without sacrificing accuracy. The duration between the appearance of the target object and the pointing response was termed response latency. The participants' pointing directions were also recorded to calculate angular pointing errors.

After the participants finished the first block of the trials (i.e., 12 trials) in the immediate room, the remote room appeared. The participants were instructed to move to the remote room and finish the second block of the trials.

2.1.4 Data analysis

For the remote perspective-taking task, we calculated the mean orientation latency, mean response latency and mean absolute angular pointing error in each type of trial. To test the global/local sensorimotor alignment effect, we conducted two-tailed paired sample *t* tests on orientation latency, response latency and absolute pointing error to see if the participants performed better in the globally/locally aligned trials than the globally/locally misaligned trials.

For null effects, we also calculated the Bayes Factor (BF_{01}) favoring the null effect over the alternative.[6]

2.2 Results

No significant results from orientation latency were produced for either the global or local sensorimotor alignment effect in any experiment (Figure S4 in supplementary materials). Thus, we are only reporting detailed results from response latency and absolute pointing error.

2.2.1 Response latency

Figure 2 shows the mean response latency for each trial type in all experiments. There was no significant difference between the globally aligned and globally misaligned conditions, $t(47) = 0.42$, $p = .674$, Cohen's $d = 0.09$, $BF_{01} = 8.11$, indicating no global sensorimotor alignment effect.

The response latency in the locally aligned condition was significantly smaller than that in the locally misaligned condition, $t(47) = 3.13$, $p = .003$, Cohen's $d = 0.64$, indicating a local sensorimotor alignment effect.

[Figure 2]

2.2.2 Absolute pointing error

Figure 3 shows the mean absolute angular pointing error for each trial type in all experiments. There was no significant difference between the globally aligned and globally misaligned conditions, $t(47) = 1.31$, $p = .197$, Cohen's $d = 0.27$, $BF_{01} = 3.88$, indicating no global sensorimotor alignment effect.

[6] The null effect is favored if the BF_{01} is larger than three, and strongly favored if the BF_{01} is larger than ten. The alternative effect is favored if the BF_{01} is smaller than 1/3, and strongly favored if the BF_{01} is smaller than 1/10 (Rouder et al., 2009). If the BF_{01} is between 1/3 and three, neither is favored.

There was no significant difference between the locally aligned and locally misaligned conditions, $t(47) = 1.74$, $p = .088$, Cohen's $d = 0.36$, $BF_{01} = 2.08$, indicating no clear evidence for a local sensorimotor alignment effect.[7]

[Figure 3]

2.3 Discussion

The results in Experiment 1 indicate a local sensorimotor alignment effect but no global sensorimotor alignment effect. These results are not consistent with the *dual means* hypothesis which stipulates that people use the means of *global-path-integration* as well as the means of *visual-re-anchoring* to update their headings relative to a remote space.

The findings of Experiment 1 may not conclusively distinguish between the *visual-re-anchoring only* hypothesis and the *conditional dual means* hypothesis, because both hypotheses can explain the null result of the global sensorimotor alignment effect. Following the *visual-re-anchoring only* hypothesis, this null result was due to the possibility that the participants could not update headings in sensorimotor global representations. However, following the *conditional dual means* hypothesis, the null global sensorimotor alignment effect might be attributed to no sensorimotor global representations produced to update. The participants could have updated headings globally when global representations were on the sensorimotor level. To test these two hypotheses, Experiments 2 and 3 were designed to activate sensorimotor global representations.

3. Experiment 2

Experiment 2 was identical to Experiment 1 except that the participants did a task of relative heading judgments of two views before the remote perspective-taking task. In particular,

[7] We also calculated correlations between response latency and absolute pointing error. There were no significant correlations in Experiments 1-3 ($r(46)=0.041$, $p=0.781$; $r(46)=-0.146$, $p=0.324$; $r(46)=-0.106$, $p=0.563$, respectively).

while standing in the immediate room, the participants in Experiment 2 were asked to face the global cardinal direction of a probed heading in the remote room (across-room relative heading judgment task, Burte & Hegarty, 2014) before the remote perspective-taking task. The across-room relative heading judgment task was used to encourage the participants to retrieve the global representations from long-term memory to the sensorimotor level as it probed global spatial relations between the two rooms. If the results showed a global sensorimotor alignment effect in addition to a local sensorimotor alignment effect, the *conditional dual means* hypothesis would be supported. If the results still did not show a global sensorimotor alignment effect but showed only a local sensorimotor alignment effect, the *visual-re-anchoring only* hypothesis would be supported.

3.1 Method

3.1.1 Participants

Forty-eight university students (24 females) with normal or corrected-to-normal vision participated to partially fulfill the requirement for an introductory psychology course.

3.1.2 Materials, design, and procedure

The materials, design, and procedure were the same as in Experiment 1 except that the relative heading judgment task was added prior to the remote perspective-taking task in the current experiment.

In the relative heading judgment task, for each trial the participants moved to face an actual view in the immediate room and then were probed with another view. The participants were asked to turn in place to face the cardinal direction of the probed view. For example, if the participants were actually facing View 14 and the probed view was View 1, then they were expected to turn 90° clockwise to face the global direction of the probed view. The participants did this task only in the first testing room, which was the actual room in the first block of the remote perspective-taking task. Table 2 shows all the trials used in the relative heading judgment task when the participants were actually in Room 2 shown in Figure 1. There were eight trials in total. Four trials were across-room trials in which the probed views were from the remote room, and the other four trials were within-room trials in which the probed views were from the immediate room. In the four trials of each type, the global angular differences between the actual and probed views were 0º, 90º (both clockwise and counterclockwise), and 180º. The within-room trials were used to ensure that the participants understood the task and conducted it correctly. In the across-room trials, the probed views were the initial views, which was consistent with the remote perspective-taking task that also used the initial views as the imagined views.

[Table 2]

For each trial of the relative heading judgment task, the participants were instructed to move and face an actual view (e.g., "Move to face the clock"). After the participants moved onto the corresponding stage, the correct object was presented. The participants were directed to the correct stage if they moved to a wrong stage. Facing the correct object, they pressed a button on a gamepad. Then the virtual environment disappeared and the participants saw only a black screen. The participants were instructed to physically turn to face the direction of a probed view (e.g., "Imagine the experimenter is facing the bin. Turn to face the same direction") and then to press a button to indicate the end of their response. The participants were instructed to respond as quickly as possible without sacrificing accuracy. We recorded the participants' response headings. The duration between the presentation of the probed view and when the participant pressed the ending button was recorded as response latency.

After the relative heading judgment task, the participants were not required to face any specific direction until the start of the remote perspective-taking task. Specifically, they were required to move to face the actual view of the first testing trial in the remote perspective-taking task.

3.2 Results

For brevity, we only present the results for the remote perspective-taking task below. The results for the relative heading judgment task can be found in the supplementary materials.

3.2.1 Response latency

As shown in Figure 2, there was no significant difference between the globally aligned and globally misaligned conditions, $t(47) = 0.44$, $p = .666$, Cohen's $d = 0.09$, $BF_{01} = 8.07$, indicating no global sensorimotor alignment effect.

The response time in the locally aligned condition was significantly smaller than that in the locally misaligned condition, $t(47) = 4.13$, $p < .001$, Cohen's $d = 0.84$, indicating a local sensorimotor alignment effect.

3.2.2 Absolute pointing error

As illustrated in Figure 3, the absolute pointing error in the globally aligned condition was significantly smaller than that in the globally misaligned condition, $t(47) = 2.62$, $p = .012$, Cohen's $d = 0.54$. This result reveals a global sensorimotor alignment effect, suggesting that the participants updated their headings in the global representations.

The responses were more accurate in the locally aligned than those in the locally misaligned conditions, $t(47) = 2.46$, $p = .018$, Cohen's $d = 0.50$, indicating a local sensorimotor alignment effect.

3.3 Discussion

The results in Experiment 2 show the global sensorimotor alignment effect in addition to the local sensorimotor alignment effect. Compared with the null global sensorimotor alignment effect in Experiment 1, this indicates that people can update their headings in the global representations to take a remote perspective if the global representations are activated to be on the sensorimotor level by across-room relative heading judgments. These findings support the *conditional dual means* hypothesis.

4. Experiment 3

In Experiment 2, the remote perspective-taking task only included the initial views as the imagined views (i.e., Views 1, 2, 9, and 10, which were the first views seen when entering the rooms). Experiment 3 was designed to replicate the global and local sensorimotor alignment effects in Experiment 2 by including all views as the imagined views. The non-initial views (i.e., views other than the initial views, such as View 4) were used as the imagined views in addition to the initial views. Previous studies examining top-down updating have shown priming effects (i.e., the advantage of mental perspective-taking across same views versus across different views in two consecutive trials) occurred only for the initial views but not for the non-initial views (Avraamides & Kelly, 2005; Lei et al., 2020). This might be because the initial views define the principal local reference directions of spatial memories (Marchette et al., 2014). When retrieving long-term spatial memories, the non-initial views might need to be transformed from the reference directions (i.e., the initial views) so that the priming effects might be smaller or diminished for imagining the non-initial views (Lei et al., 2020). However, as the global and local sensorimotor effects in Experiment 2 occurred due to the interference between two representations, one from bottom-up updating and the other from top-down updating, such interference might occur for all views, leading to the global and local sensorimotor effects for all views.

4.1 Method

4.1.1 Participants

Thirty-two university students (16 females) with normal or corrected-to-normal vision participated to partially fulfill the requirement for an introductory psychology course.

The number of participants decreased in Experiment 3 compared with Experiments 1 and 2. The reason was that the number of trials in the remote perspective-taking task increased in Experiment 3. In Experiments 1 and 2, the task had 24 trials with the imagined views being the initial views which were two out of the eight views in each room. In Experiment 3, the imagined views included all eight views in each room and thus the number of trials increased to 96. In Experiment 2, the effect size of the global sensorimotor alignment effect was medium sized (Cohen's $d = 0.54$). Assuming the effect size in the current Experiment 3 increased to be large sized (Cohen's $d = 0.80$), the power of using 32 participants was 0.87 at the 0.05 level (for two-tailed paired t test, see the Matlab code for the power analysis at https://doi.org/10.7939/r3-vm8t-xy36)

4.1.2 Materials, design, and procedure

The materials, design, and procedure in Experiment 3 were the same as in Experiment 2 except the trials in the remote perspective-taking task. In each block of the remote perspective-taking task, there were 48 trials, with 12 trials in each of the four types of trials (i.e., globally aligned/misaligned, locally aligned/misaligned). For each type of trial, there were four pairs of the actual and imagined views. Table 3 shows all pairs of the actual and imagined views when the actual views were in Room 2 shown in Figure 1. As in previous experiments, in both globally (locally) aligned and globally (locally) misaligned trials, the local (global) angular difference between the actual and imagined views was 90° to remove any confound from the local (global) sensorimotor alignment effect. As in previous experiments, three trials were generated for each view pair with three different targets. A sample trial list can be found at https://doi.org/10.7939/r3-ch2r-xr18.

[Table 3]

4.2 Results

Below are the results for the remote perspective-taking task. The results for the relative heading judgment task can be found in the supplementary materials.

4.2.1 Response latency

As shown in Figure 2, there was no significant difference between the globally aligned and globally misaligned conditions, $t(31) = 0.50$, $p = .621$, Cohen's $d = 0.12$, $BF_{01} = 6.47$, indicating no global sensorimotor alignment effect.

The response latency in the locally aligned condition was significantly smaller than that in the locally misaligned condition, $t(31) = 4.71$, $p < .001$, Cohen's $d = 1.18$, indicating a local sensorimotor alignment effect.

4.2.2 Absolute pointing error

As shown in Figure 3, the absolute pointing error in the globally aligned condition was significantly smaller than that in the globally misaligned condition, $t(31) = 3.24$, $p = .003$, Cohen's $d = 0.81$, indicating a global sensorimotor alignment effect.

The absolute pointing error in the locally aligned condition was significantly smaller than that in the locally misaligned condition, $t(31) = 2.72$, $p = .011$, Cohen's $d = 0.68$, indicating a local sensorimotor alignment effect.

4.3 Discussion

Experiment 3 showed the global sensorimotor alignment effect in addition to the local sensorimotor alignment effect in remote perspective-taking, which replicated the results in Experiment 2. These results again supported the *conditional dual means* hypothesis.

5. General Discussion

The current study examined global and local sensorimotor alignment effects to infer how people rely on path integration and piloting to update self-location in familiar multiscale spaces with the support of multiscale spatial memories. The findings showed a local sensorimotor alignment effect based on visual similarity between the immediate and remote environments and a global sensorimotor alignment effect based on self-motion cues in across-boundary navigation. However, the global sensorimotor alignment effect required activated sensorimotor global representations. These findings favor the *conditional dual means* hypothesis.

To the best of our knowledge, the current study is the first one demonstrating a global sensorimotor alignment effect in remote perspective-taking in the literature. Furthermore, the current study showed that the presence of global representations in long-term memory is not sufficient for a global sensorimotor alignment effect; rather global spatial representations on the sensorimotor level are required. Only when the global representations were activated by the across-room relative heading judgments prior to the remote perspective-taking task, the participants updated self-location in both global and local representations. Otherwise, the participants only updated local self-localization representations.

The *conditional dual means* hypothesis can explain why no global sensorimotor alignment effect in remote perspective-taking was reported in Kelly et al. (2007). In their study, the participants moved along a simple path from a learning space to a testing space and could point accurately to the learning space from the testing position. They should have learned global spatial relations between the learning and testing spaces. We speculate that pointing to the learning space from the testing position might not be strong enough to activate the global representations of objects' locations in the learning space on the sensorimotor level. The participants in Kelly et al. might have used only the homing vector (from the end to the origin of the path) instead of using the global representations of spatial relations between the objects in the learning space and their actual headings in the testing space to point to the learning space. Thus no global representations might have been retrieved to the sensorimotor level to support updating on the global level, leading to no global sensorimotor alignment effect in their study.

Expanding the *conditional dual means* hypothesis, we propose a model (Figure 4) to sketch the interaction between spatial navigation and spatial memory. This model emphasizes the role of sensorimotor representations in navigation in a familiar multiscale environment with the support of multiscale spatial memories, and in forming multiscale spatial memories from across-boundary navigation in a novel multiscale environment (Lei et al., 2020).

[Figure 4]

In navigation, path integration relies on self-motion cues to update one's self-location continuously (Etienne & Jeffery, 2004; Loomis et al., 1999; Mittelstaedt & Mittelstaedt, 1980; Wang, 2017; Yamamoto & Shelton, 2005). Although path integration accesses sensorimotor representations in working memory to update continuously, we speculate that path integration may not directly access spatial representations in long-term memory (Shrager et al., 2008). This speculation is based on the assumption that accessing spatial representations in long-term memory is much slower than in working memory. It is not efficient to continuously access spatial representations in long-term memory to support continuous updating. Thus, whether path integration updates globally or locally depends on the nature of the sensorimotor representations. If sensorimotor representations are global, path integration updates globally; if sensorimotor representations are local, path integration updates locally.

In contrast, previous studies have indicated that piloting, which mainly relies on visual cues and updates intermittently, not only updates self-localization representations on the sensorimotor level (Riecke & McNamara, 2017) but also can change sensorimotor representations by directly retrieving long-term memory (Etienne et al., 2004; Sholl et al., 2006). It pinpoints one's self-location in mental maps in long-term memory and brings this self-localization representation from long-term memory to the sensorimotor level whenever the visual cues are inconsistent with the sensorimotor representations. For example, when people are distracted from path integration (e.g., engaging in a conversation with others), sensorimotor representations are not updated continuously. People may retrieve spatial information associated with some visual cues from long-term memory to regenerate sensorimotor representations to resume path integration. This refers to resetting path integration by piloting (Etienne et al., 2004; Zhang & Mou, 2017). Importantly, piloting can also switch sensorimotor representations from local to global representations or vice versa. For example, the participants in Sholl et al. (2006) switched sensorimotor representations from local to global representations by looking at the campus outside the window of the lab.

Control processes based on instructions and intentions independent of the navigation cues (self-motion or piloting cues) may also retrieve long-term memory to the sensorimotor level. When global relations are goal-relevant, global representations can be brought to the sensorimotor level. This is supported by the current findings that relative heading judgments between headings in the two rooms could retrieve global representations on the sensorimotor level. Other studies have also shown that instructions to understand and attend to global relations can elicit sensorimotor global representations (Burte & Hegarty, 2014; Röhrich et al., 2014; Shine et al., 2016; Sholl et al., 2006). Therefore, we speculate that control processes can also affect whether sensorimotor representations are local or global.

Importantly, the findings of local but no global sensorimotor alignment effects in Experiment 1 of the current study and in the previous studies (e.g., Kelly et al., 2007) suggest that even when multiscale spatial memories are available, sensorimotor representations may be primarily based on local spatial representations during navigation. This may be because (a) people usually have rich piloting cues for the immediate environment; (b) people primarily engage with the immediate environment (e.g., Wang, 2004), so the control processes also focus on sensorimotor local representations. The human sensorimotor system is to guide coordination in movement (Saltzman, 1979). Local representations rather than global representations are closely relevant to the movement in the immediate space. However, these primary sensorimotor local representations can be overridden by the piloting cues indicating the global environments, or by the control processes to focus on the global environments as indicated by the findings of the global sensorimotor alignment effects in Experiments 2 and 3 of the current study. Thus, the current model can explain the findings regarding the global sensorimotor alignment effects in the current study and in the previous studies (e.g., Kelly et al., 2007).

The current model (Figure 4) is also applicable to forming multiscale spatial representations in long-term memory through across-boundary navigation. It is a long-established theoretical speculation that path integration is critical to developing global spatial memories, as piloting cues which indicate the global spatial relations are usually blocked in a large-scale environment (Gallistel, 1990; Jacobs & Schenk, 2003; Loomis et al., 1999;

Meilinger, 2008). Studies have also shown that people can form global spatial representations by across-boundary navigation without piloting cues of the global spatial relations (Holmes et al., 2018; Lei et al., 2020; Shine et al., 2016; Yamamoto & Shelton, 2005). However, developing global memories across boundaries requires preconditions including explicit instructions to encode global relations, sufficient learning time or some prior global learning (Han & Becker, 2014; Lei et al., 2020; Shine et al., 2016). Expanding these ideas and findings, our current model underscores the importance of sensorimotor global representations in developing global spatial memories by path integration through across-boundary navigation. We conjecture that with the presence of sensorimotor global representations (under the preconditions of explicit instructions, sufficient learning time, or some prior global learning), path integration integrates the current space into the existing global representations, and thus the existing global framework scaffolds the newly-acquired spatial information of local spaces and integrates them into larger/newer global representations in long-term memory (e.g., Jacobs & Schenk, 2003; Kelly & McNamara, 2010). By contrast, without the presence of sensorimotor global representations, path integration does not update headings in global spatial representations, so no development of larger/newer global representations occurs (e.g., Marchette et al., 2014).

Distinguishing spatial representations on the sensorimotor level from those in long-term memory is consistent with the theory claiming both online and offline updating systems (Easton & Sholl, 1995; Hartley & Burgess, 2005; Kelly et al., 2007; Mou et al., 2004; Waller & Hodgson, 2006; Wang & Spelke, 2000). People may have an offline system to represent interobject spatial relations in long-term memory. People also have an online system to represent self-to-objects spatial relations in their sensorimotor system or working memory. While the offline system (long-term memory) has both local and global spatial representations, the online system (sensorimotor representations) primarily has local spatial representations.

The current finding of the global sensorimotor alignment effect in remote perspective-taking occurred under the condition of global spatial memories acquired from substantial across-boundary navigation (about eight times moving back and forth between the two rooms in the learning phase of the current study and Lei et al. (2020)). It could be because the participants in these two studies physically turned but used a moving device to visually translate in the virtual environment, which diminished the role of self-motion cues and might have impaired spatial updating during navigation (Chance et al., 1998; Ruddle et al., 2011). It is not clear whether the current findings can be generalized to a situation in which people move between spaces only once but with full self-motion cues as in Kelly et al. (2007). On the one hand, from one-time navigation across spaces, people might not be adept at developing or integrating global representations of headings across spaces in long-term memory (Starrett et al., 2019). On the other hand, one-time navigation across spaces might be sufficient to develop global spatial memory (Ishikawa & Montello, 2006) but the global spatial representations might not be activated on the sensorimotor level to produce a global sensorimotor alignment effect (Kelly et al., 2007). Future studies are needed to test updating self-localization representations in a remote space when there are few experiences to navigate between the remote and immediate spaces.

In the current study, the experimental setup and the learning procedure were the same as those in Lei et al. (2020). Lei et al. (see also the experiment in the supplementary materials) showed a significant global priming effect in addition to a local priming effect. The consistent findings of the global priming effect and the global sensorimotor alignment effect provide converging evidence for the global spatial representations. However, sensorimotor alignment effects are caused by the congruency between the representation of the actual heading from bottom-up updating and the representation of the imagined heading from top-down updating. By contrast, priming effects are caused by two representations of imagined perspectives for two consecutive trials, both from top-down updating.

In conclusion, the current findings indicate that after developing multiscale spatial representations of remote and immediate spaces, people can rely on both piloting and path integration in bottom-up updating of their headings. People update headings in the local representations of the remote space based on local structures/visual similarities between immediate and remote spaces, and update headings in global representations of both spaces based on self-motion cues available in across-boundary navigation. Nevertheless, updating headings globally might require activating sensorimotor global representations.

Acknowledgments

This work was funded by the NSERC, Canada to Weimin Mou. We thank Subekshya Adhikari, Jarlo Alganion, Aila Jamali, Dhriti Mehta, and Zehua Xia for their contribution to data collection.

References

Attneave, F., & Farrar, P. (1977). The visual world behind the head. *The American journal of psychology*, 549-563.

Avraamides, M. N. (2003). Spatial updating of environments described in texts. *Cognitive Psychology*, 47(4), 402-431.

Avraamides, M. N., & Kelly, J. W. (2005). Imagined perspective-changing within and across novel environments. In C. Freksa, B. Nebel, M. Knauff, & B. Krieg-Brückner (Eds.), *Lecture notes in artificial intelligence: Spatial cognition IV. Reasoning, action, interaction* (pp. 245–258). Berlin: Springer.

Avraamides, M. N., & Kelly, J. W. (2008). Multiple systems of spatial memory and action. *Cognitive processing*, *9*(2), 93-106.

Brockmole, J. R., & Wang, R. F. (2002). Switching between environmental representations in memory. *Cognition*, *83*(3), 295-316.

Burte, H., & Hegarty, M. (2014). Alignment effects and allocentric-headings within a relative heading task. In C. Freksa, B. Nebel, M. Hegarty & T. Barkowsky (Eds.), *Spatial Cognition IX* (pp. 46-61). Cham: Springer.

Chance, S. S., Gaunet, F., Beall, A. C., & Loomis, J. M. (1998). Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence*, *7*(2), 168-178.

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition*, *23*(2), 149-178.

Cheng, K., & Spetch, M. L. (1998). Mechanisms of landmark use in mammals and birds. In S. Healy (Ed.), *Spatial representation in animals* (pp. 1–17). New York, NY: Oxford University Press.

Doeller, C. F., & Burgess, N. (2008). Distinct error-correcting and incidental learning of location relative to landmarks and boundaries. *Proceedings of the National Academy of Sciences*, *105*(15), 5909-5914.

Easton, R. D., & Sholl, M. J. (1995). Object-array structure, frames of reference, and retrieval of spatial knowledge. *Journal of experimental psychology: learning, memory, and cognition*, *21*(2), 483.

Etienne, A. S., & Jeffery, K. J. (2004). Path integration in mammals. *Hippocampus, 14*(2)*,* 180-192.

Etienne, A. S., Maurer, R., Boulens, V., Levy, A., & Rowe, T. (2004). Resetting the path integrator: a basic condition for route-based navigation. *Journal of Experimental Biology*, *207*(9), 1491-1508.

Foo, P., Warren, W. H., Duchon, A., & Tarr, M. J. (2005). Do humans integrate routes into a cognitive map? Map-versus landmark-based navigation of novel shortcuts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(2), 195.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: The MIT Press.

Han, X., & Becker, S. (2014). One spatial map or many? Spatial coding of connected environments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 511-531.

Hartley, T., & Burgess, N. (2005). Complementary memory systems: competition, cooperation and compensation. *Trends in Neurosciences*, *28*(4), 169-170.

He, Q., & McNamara, T. P. (2018). Spatial updating strategy affects the reference frame in path integration. *Psychonomic bulletin & review*, *25*(3), 1073-1079.

Holmes, C. A., Newcombe, N. S., & Shipley, T. F. (2018). Move to learn: Integrating spatial information from multiple viewpoints. *Cognition*, *178*, 7-25.

Ishikawa, T., & Montello, D. R. (2006). Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive psychology*, *52*(2), 93-129.

Jacobs, L. F., & Schenk, F. (2003). Unpacking the cognitive map: the parallel map theory of hippocampal function. *Psychological review*, *110*(2), 285.

Kelly, J. W., Avraamides, M. N., & Loomis, J. M. (2007). Sensorimotor alignment effects in the learning environment and in novel environments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(6), 1092-1107.

Kelly, J. W., & McNamara, T. P. (2010). Reference frames during the acquisition and development of spatial memories. *Cognition*, *116*(3), 409-420.

Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science, 9*(4), 293-298.

Liu, C., & Xiao, C. (2018). Dual Systems for Spatial Updating in Immediate and Retrieved Environments: Evidence from Bias Analysis. *Frontiers in Psychology*, *9*, 85.

Lei, X., Mou, W., & Zhang, L. (2020). Developing global spatial representations through across-boundary navigation. *Journal of experimental psychology. Learning, memory, and cognition*, *46*(1), 1.

Loomis, J.M., Klatzky, R.L., Avraamides, M., Lippa Y., & Golledge, R.G. (2007). Functional Equivalence of Spatial Images Produced by Perception and Spatial Language. In F. Mast, & L. Jäncke (Eds.), *Spatial processing in navigation, imagery and perception* (pp. 29-48). Boston, MA: Springer.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Philbeck, J. W. (1999). Human navigation by path integration. In R. G. Golledge (Ed.), *Wayfinding: Cognitive mapping and other spatial processes* (pp. 125–151). Baltimore, MD: Johns Hopkins University Press.

Marchette, S. A., Vass, L. K., Ryan, J., & Epstein, R. A. (2014). Anchoring the neural compass: coding of local spatial reference frames in human medial parietal lobe. *Nature Neuroscience, 17*(11), 1598-1606.

May, M. (2007). Imaginal repositioning in everyday environments: Effects of testing method and setting. *Psychological Research*, *71*(3), 277-287.

May, M., & Klatzky, R. L. (2000). Path integration while ignoring irrelevant movement. *Journal of experimental psychology: learning, memory, and cognition*, *26*(1), 169.

McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., & Moser, M. B. (2006). Path integration and the neural basis of the'cognitive map'. *Nature Reviews Neuroscience*, *7*(8), 663.

Meilinger, T. (2008). The network of reference frames theory: A synthesis of graphs and cognitive maps. In C. Freksa, N. S. Newcombe, P. Gärdenfors & S. Wölfl (Eds.), *Spatial Cognition VI. Learning, Reasoning, and Talking about Space* (pp. 344-360). Berlin: Springer.

Mittelstaedt, M. L., & Mittelstaedt, H. (1980). Homing by path integration in a mammal. *Naturwissenschaften, 67*(11), 566-567.

Mou, W., McNamara, T. P., Valiquette, C. M., & Rump, B. (2004). Allocentric and egocentric updating of spatial memories. *Journal of experimental psychology: Learning, Memory, and Cognition*, *30*(1), 142.

Mou, W., & Wang, L. (2015). Piloting and path integration within and across boundaries. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(1), 220-234.

Montello, D. R. (1993). Scale and multiple psychologies of space. In A. U. Frank & I. Campari (Eds.), *Spatial information theory* (pp. 312–321). Berlin, Germany: Springer.

Philbeck, J. W., & O'Leary, S. (2005). Remembered landmarks enhance the precision of path integration. *Psicológica*, *26*, 7–24.

Riecke, B. E., & McNamara, T. P. (2017). Where you are affects what you can easily imagine: Environmental geometry elicits sensorimotor interference in remote perspective taking. *Cognition*, *169*, 1-14.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1157.

Rieser, J. J., Guth, D. A., & Hill, E. W. (1986). Sensitivity to perspective structure while walking without vision. *Perception*, *15*(2), 173-188.

Röhrich, W. G., Hardiess, G., & Mallot, H. A. (2014). View-based organization and interplay of spatial working and long-term memories. *PloS one*, *9*(11), e112793.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic bulletin & review*, *16*(2), 225-237.

Ruddle, R. A., Volkova, E., & Bülthoff, H. H. (2011). Walking improves your cognitive map in environments that are large-scale and large in extent. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *18*(2), 1-20.

Saltzman, E. (1979). Levels of sensorimotor representation. *Journal of Mathematical Psychology*, *20*(2), 91-163.

Shelton, A. L., & Marchette, S. A. (2010). Where do you think you are? Effects of conceptual current position on spatial memory performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 686.

Shine, J. P., Valdés-Herrera, J. P., Hegarty, M., & Wolbers, T. (2016). The human retrosplenial cortex and thalamus code head direction in a global reference frame. *The Journal of Neuroscience, 36*(24), 6371-6381.

Sholl, M. J., Kenny, R. J., & DellaPorta, K. A. (2006). Allocentric-heading recall and its relation to self-reported sense-of-direction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(3), 516.

Shrager, Y., Kirwan, C. B., & Squire, L. R. (2008). Neural basis of the cognitive map: path integration does not require hippocampus or entorhinal cortex. *Proceedings of the National Academy of Sciences*, *105*(33), 12034-12038.

Starrett, M. J., Stokes, J. D., Huffman, D. J., Ferrer, E., & Ekstrom, A. D. (2019). Learning-dependent evolution of spatial representations in large-scale virtual environments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(3), 497.

Waller, D., & Hodgson, E. (2006). Transient and enduring spatial representations under disorientation and self-rotation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 867.

Waller, D., Montello, D. R., Richardson, A. E., & Hegarty, M. (2002). Orientation specificity and spatial updating of memories for layouts. *Journal of experimental psychology: Learning, Memory, and Cognition*, *28*(6), 1051.

Wang, R. F. (2004). Between reality and imagination: When is spatial updating automatic?. *Perception & psychophysics*, *66*(1), 68-76.

Wang, R. F. (2016). Building a cognitive map by assembling multiple path integration systems. *Psychonomic Bulletin & Review, 23*(3), 692-702.

Wang, R. F. (2017). Spatial updating and common misinterpretations of spatial reference frames. *Spatial Cognition & Computation*, *17*(3), 222-249.

Wang, R. F., & Brockmole, J. R. (2003a). Simultaneous spatial updating in nested environments. *Psychonomic bulletin & review*, *10*(4), 981-986.

Wang, R. F., & Brockmole, J. R. (2003b). Human navigation in nested environments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(3), 398-404.

Wang, R. F., & Spelke, E. S. (2000). Updating egocentric representations in human navigation. *Cognition*, *77*(3), 215-250.

Wehner, R., Michel, B., & Antonsen, P. (1996). Visual navigation in insects: coupling of egocentric and geocentric information. *Journal of Experimental Biology*, *199*(1), 129-140.

Wiener, J. M., Berthoz, A., & Wolbers, T. (2011). Dissociable cognitive mechanisms underlying human path integration. *Experimental brain research*, *208*(1), 61-71.

Yamamoto, N., & Shelton, A. L. (2005). Visual and proprioceptive representations in spatial memory. *Memory & Cognition, 33*(1), 140-150.

Zhang, L., & Mou, W. (2017). Piloting systems reset path integration systems during position estimation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(3), 472.

Figures

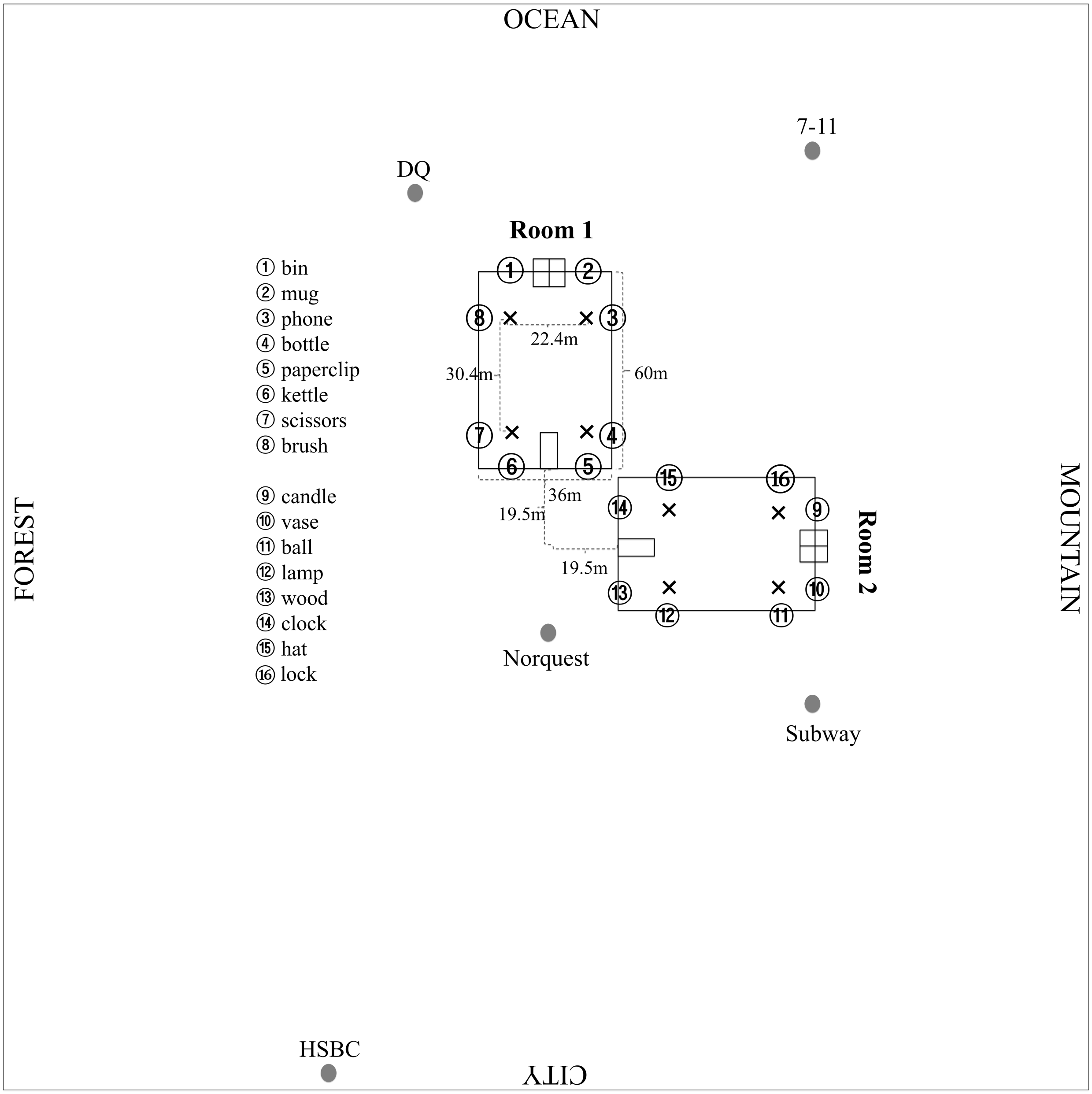

Figure 1. Top view of schematic experimental setup. The four distal orientation cues are infinitely distant. The dots are five buildings. In Rooms 1 and 2, the numbers are views, and the crosses are standing stages to see the views. The specific objects for individual views are listed. The room size, distances between the standing stages and distances between the center of the environment and the rooms are illustrated. The distances between the center of the environment

and the buildings are scaled down by 3.2 times to fit in the figure. The rooms and the buildings were not simultaneously presented. When the buildings were presented, only the doorways of the rooms were presented. When the rooms were presented, the buildings disappeared.

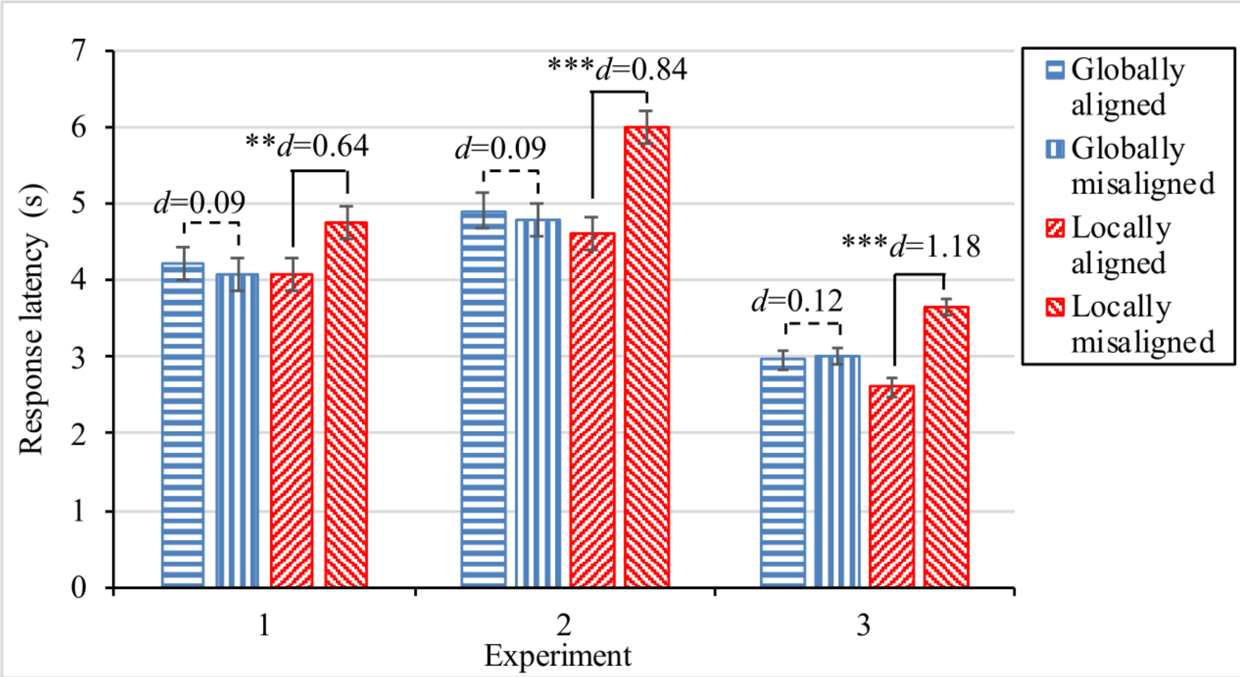

Figure 2. *The mean response latency for four types of trials in the remote perspective-taking task in all experiments. Error bars represent ±1 SE removing the variance from individual differences. The solid lines are for significant comparisons. The dashed lines are for insignificant comparisons. Cohen's d values are listed (** p< .01; *** p< .001).*

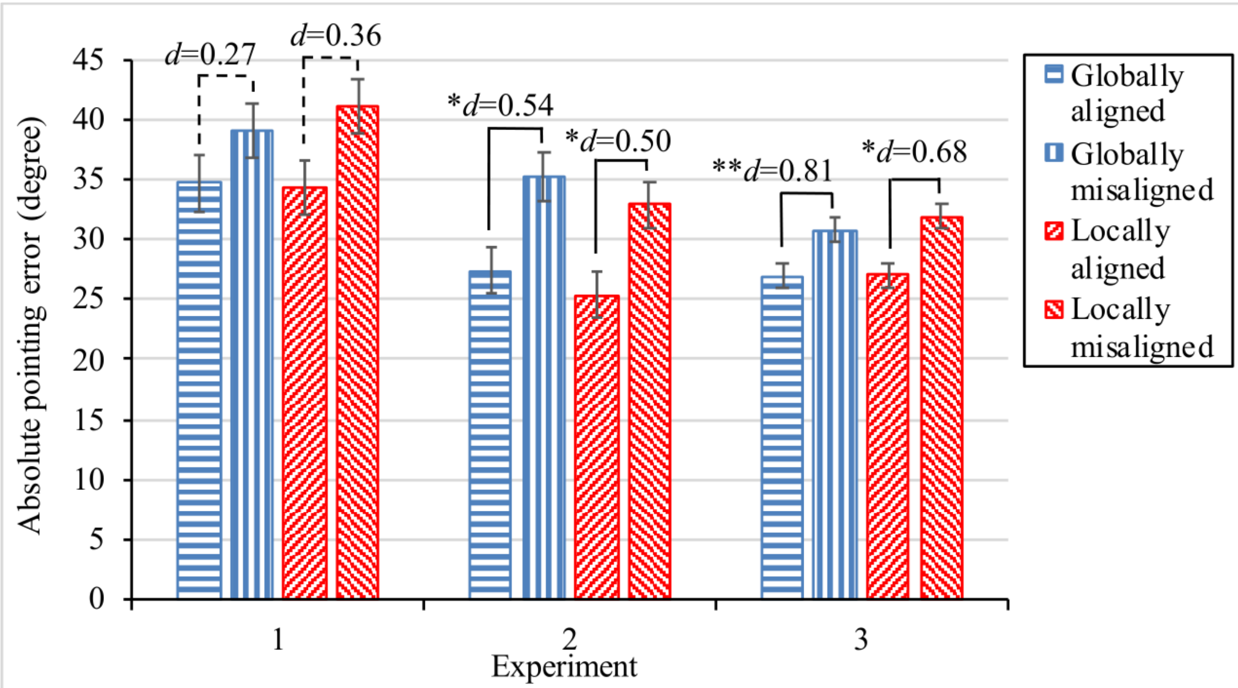

Figure 3. The mean absolute angular pointing error for four types of trials in the remote perspective taking task in all experiments. Error bars represent ±1 SE removing the variance from individual differences. The solid lines are for significant comparisons. The dashed lines are for insignificant comparisons. Cohen's d values are listed (p< .05; ** p< .01).*

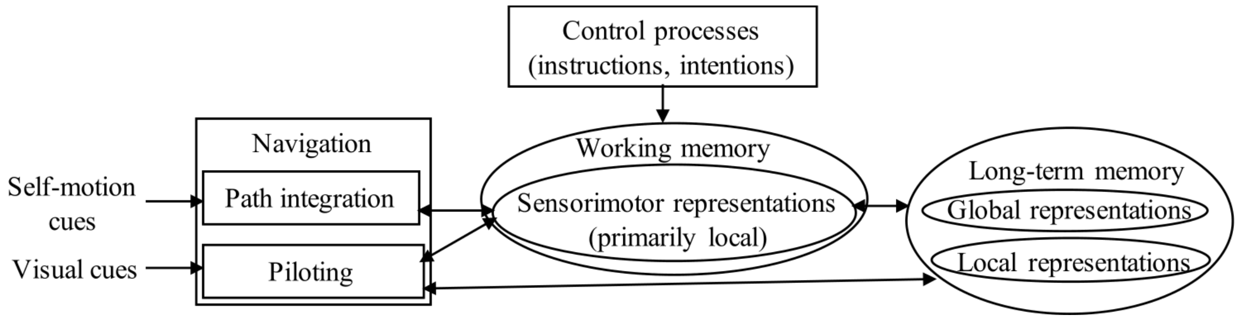

Figure 4. *Schematic diagram to stipulate how path integration and piloting in navigation interact with spatial representations in long-term memory through sensorimotor representations. Path integration mainly relies on self-motion cues and accesses sensorimotor representations to interact with spatial representations in long-term memory. Piloting mainly relies on visual cues and interacts with spatial representations in long-term memory either directly or through sensorimotor representations. Control processes (instructions, intentions) can manipulate sensorimotor representations to be generated based on local or global spatial representations. Circles are for representations in memory, while rectangles are for processes.*

Tables

Table 1

Sample views for four trial types in the remote perspective-taking task in Experiments 1 and 2. The actual view was the view that the participants were actually facing. The imagined view was the view that the participants imagined facing. The view numbers refer to Figure 1.

Trial type	Actual view	Imagined view
Globally aligned	15	2
Globally misaligned	12	2
Locally aligned	10	1
Locally misaligned	13	1

Table 2

Views in the relative heading judgment task in Experiments 2 and 3. The actual view was the view that the participants were actually facing. The probed view was the view that the participants were asked to turn and face its direction. The view numbers refer to Figure 1.

	Actual view	Probed view	Angular difference
Across-room	16	2	0°
	14	1	90°
	11	2	180°
	9	1	90°
Within-room	16	15	0°
	9	12	90°
	11	16	180°
	14	11	90°

Table 3

Sample views for four trial types in the remote perspective-taking task in Experiment 3. The actual view was the view that the participants were actually facing. The imagined view was the view that the participants imagined facing. The view numbers refer to Figure 1.

Trial type	Actual view	Imagined view
Globally aligned	15	2
	9	4
	11	6
	13	8
Globally misaligned	12	2
	14	4
	16	6
	10	8
Locally aligned	10	1
	12	3
	14	5
	16	7
Locally misaligned	13	1
	15	3
	9	5
	11	7

Supplementary Materials

1. Replication of Lei et al. (2019)

This experiment was designed to ensure that the experimental setup and the learning procedure used in Experiments 1-3, adopted from Lei et al. (2019), could lead to global and local representations. We intended to replicate the global (and local) priming effect to demonstrate the global (and local) representations, using the priming trials based on the remote perspective-taking trials in Experiments 1-3 of the current study.

1.1 Method

1.1.1 Participants

Forty-eight university students (24 females) with normal or corrected-to-normal vision participated to partially fulfill the requirement for an introductory psychology course. This number of participants was the same as in Lei et al. (2019) and in Experiments 1 and 2 of the current study.

1.1.2 Materials and design

The experimental setup was the same as described in Experiments 1, 2 and 3 of the current study (Figure 1). After the learning phase, the participants were taken into another experimental room which looked different from the learning rooms in the virtual environment, and they did a priming task on the computer. In the priming task, the participants used a joystick (Logitech Extreme 3D Pro, Newark, CA) to judge the relative direction to a target from an imagined perspective.

The trials in the priming task were created based on the trials used in the remote perspective-taking task of the current study. In the remote perspective-taking task, every trial asked the participants to have an actual perspective (e.g., “Move to face the hat”) and then required pointing to a target from an imagined perspective (e.g., “Imagine you are facing the mug", "point to the kettle"). The relationships between actual and imagined perspectives were manipulated to be globally/locally aligned/misaligned. To keep these relationships in the priming task, we transformed one trial in the remote perspective-taking task into one pair of two consecutive trials in the priming task, so that the relationship between actual and imagined perspectives in one remote perspective-taking trial could be transformed into the relationship between imagined perspectives in one pair of two consecutive priming trials. In each pair of two consecutive priming trials, the former trial required pointing to a target from the original actual perspective in the remote perspective-taking trial (e.g., "Imagine you are facing the hat", "point to the ball"); the latter trial required pointing to a target from the original imagined perspective (e.g., "Imagine you are facing the mug", "point to the scissors"). Same as the criterion to choose a target in Experiments 1, 2 and 3, a target was randomly chosen from all objects in the same room except for the viewing object and the object on their exact back (e.g., when the participants imagined facing View 1, View 6 was on their exact back). Thus, for each pair of priming trials, the first trial was used to prime the second trial. The second trial was the one that got primed globally/locally and reflected the four trial types used in the remote perspective-taking task (i.e., globally aligned, globally misaligned, locally aligned, and locally misaligned).

The priming trials were in two blocks. The first block was generated based on the 24 trials in the remote perspective-taking task in Experiments 1 and 2, which only used the initial views as the imagined views (Table 1). The second block was generated based on the 96 trials in the remote perspective-taking task in Experiment 3, which used all views as the imagined views (Table 3). Thus, the first block had 48 priming trials and the second block had 192 priming trials (see trial details at https://doi.org/10.7939/r3-2c81-ja55).

1.2 Procedure

The learning phase in the immersive virtual environment (i.e., learning buildings and objects) was the same as in Experiments 1, 2, and 3 of the current study and also in Lei et al. (2019).

After the learning phase, the participants were led to another experimental room. They sat in a chair and did the priming task on the computer. In each trial, a sentence with white words was presented on the black screen to instruct an imagined perspective (e.g., "Imagine you are facing the hat"). The participants clicked a trigger on the joystick if they took the imagined perspective. The duration between the presentation of the imagined perspective and clicking the trigger was recorded as orientation latency. After the participants clicked the trigger, another sentence was presented to instruct a target (e.g., "point to the ball"). The participants were required to respond as quickly as possible without sacrificing accuracy. The duration between the presentation of the target and the pointing response was recorded as response latency. The participants' pointing directions were also recorded to calculate angular pointing errors. The participants were given a chance for a break when they finished half of the priming trials (i.e., 120 priming trials), and they informed the experimenter to continue the task when they were ready after the break.

1.3 Data analysis

In each pair of the priming trials, the second trial was globally/locally primed by the first trial. The performance in the second trial of each pair could reflect the influence of global/local priming. Thus, the second trial in each pair of two consecutive priming trials was used in data analysis. In addition, since the initial views were used in Lei et al. (2019) to demonstrate the global and local priming effects (see also Avraamides & Kelly, 2005), we used the trials at the initial views in data analysis. Thus, in the first block, since the imagined headings were always from the initial views, all the second trials in the pair of two consecutive priming trials were used; in the second block, only the second trials in the pair that had the imagined perspectives from the initial views were used in data analysis.

We calculated the mean response latency, mean absolute pointing error, and mean orientation latency in the four trial types (i.e., globally/locally aligned/misaligned). To test the global/local priming effects, we did repeated measures ANOVA with two within-subject factors, the block (i.e., first block versus second block) and the alignment (i.e., globally aligned versus globally misaligned when testing the global priming effect; locally aligned versus locally misaligned when testing the local priming effect).

1.4 Results

1.4.1 Response latency

Mean response latency for four types of trials in two blocks was plotted in Figure S1. For the global priming effect, the main effect of block was significant, $F(1, 47) = 15.28$, $p < .001$, $\eta_p^2 = 0.25$, which showed that the participants responded faster in the second block than in the first block. Importantly, the main effect of alignment was significant, $F(1, 47) = 4.43$, $p = .041$, $\eta_p^2 = 0.09$. The interaction between block and alignment was not significant, $F(1, 47) = 0.36$, $p = .550$, $\eta_p^2 = 0.01$. These results indicated the global priming effects in both the first and the second blocks.

For the local priming effect, the main effect of block was significant, $F(1, 47) = 4.10$, $p = .049$, $\eta_p^2 = 0.08$, indicating faster responses in the second block than in the first block. The main effect of alignment was significant, $F(1, 47) = 8.95$, $p = .004$, $\eta_p^2 = 0.16$. The interaction between block and alignment was not significant, $F(1, 47) = 0.09$, $p = .760$, $\eta_p^2 = 0.002$. These results indicated the local priming effects in both the first and the second blocks.

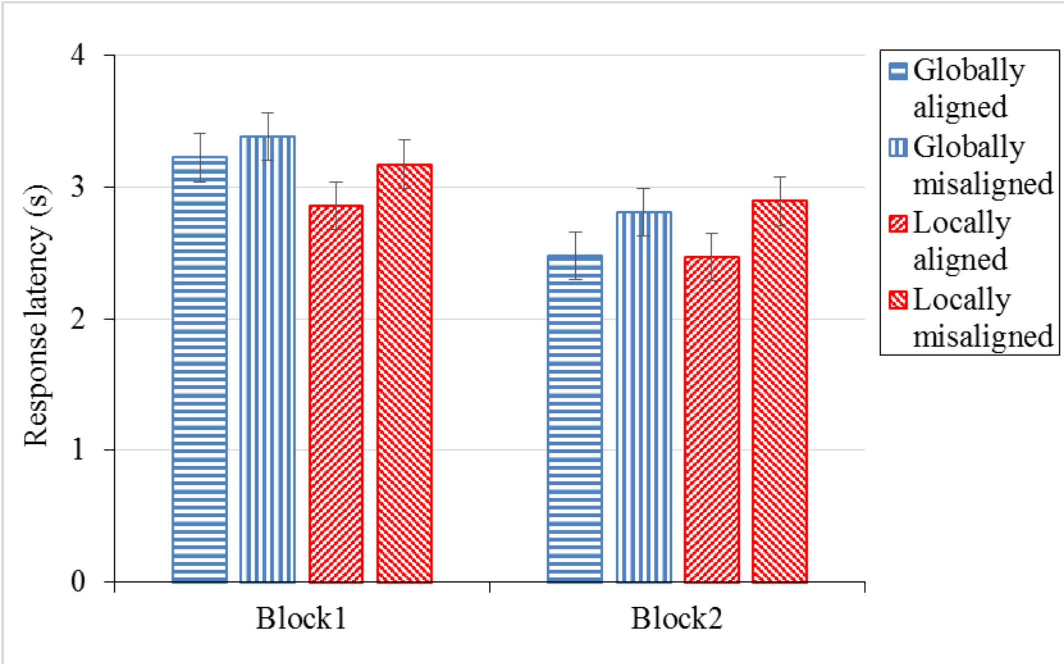

Figure S1. The mean response latency for four types of trials in two blocks of the priming task in the replication experiment. Error bars represent ±1 SE removing the variance from individual differences.

1.4.2 Absolute pointing error

Mean absolute pointing error for four types of trials in two blocks was plotted in Figure S2. For the global priming effect, none of the interaction, the main effect of block or the main effect of alignment were significant, Fs(1, 47) ≤ 1.27, ps ≥ .266, η_p^2s ≤ 0.03.

For the local priming effect, none of the interaction, the main effect of block or the main effect of alignment were significant, Fs(1, 47) ≤ 2.08, ps ≥ .156, η_p^2s ≤ 0.04.

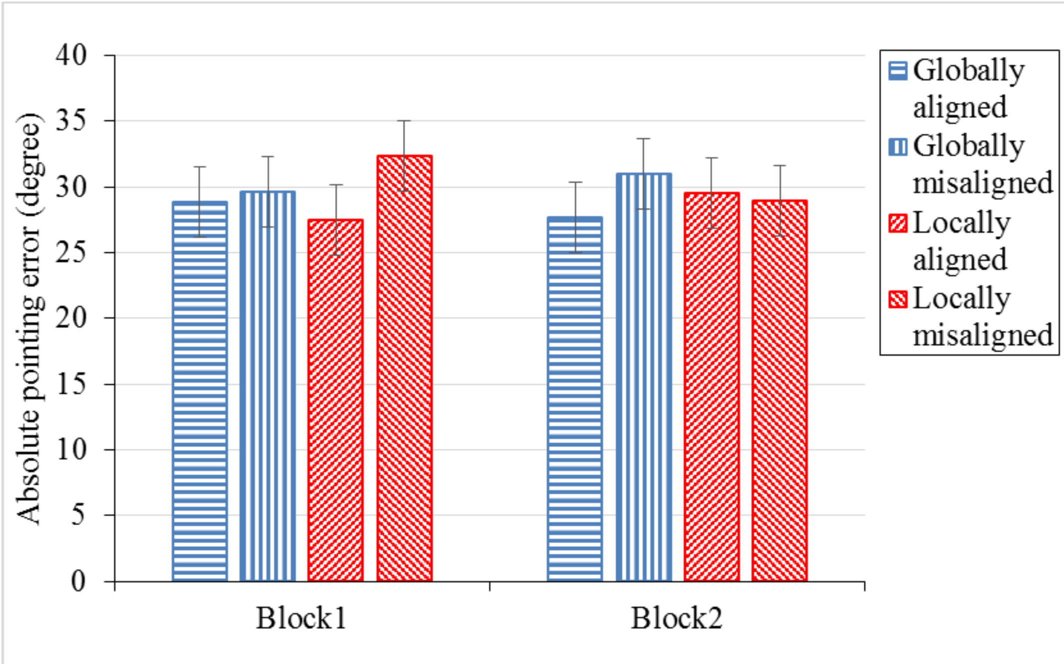

Figure S2. The mean absolute pointing error for four types of trials in two blocks of the priming task in the replication experiment. Error bars represent ±1 SE removing the variance from individual differences.

1.4.3 Orientation latency

Mean orientation latency for four types of trials in two blocks was plotted in Figure S3. For the global priming effect, the main effect of block was not significant, $F(1, 47) = 2.81$, $p = .100$, $\eta_p^2 = 0.06$. The main effect of alignment was not significant, $F(1, 47) = 1.87$, $p = .178$, $\eta_p^2 = 0.04$. The interaction between block and alignment was significant, $F(1, 47) = 9.32$, $p = .004$, $\eta_p^2 = 0.17$. Paired sample *t* tests showed that in the first block, there was no significant difference between globally aligned and globally misaligned trials, $t(47) = 0.92$, $p = .363$, Cohen's $d = 0.19$, $BF_{01} = 5.87$; in the second block, the orientation latency was faster in the globally aligned trials than the globally misaligned trials, $t(47) = 2.96$, $p = .005$, Cohen's $d = 0.60$. These results indicated the global priming effect in the second block but not in the first block.

For the local priming effect, none of the interaction, the main effect of block or the main effect of alignment were significant, $Fs(1, 47) \leq 1.93$, $ps \geq .171$, $\eta_p^2 s \leq 0.04$.

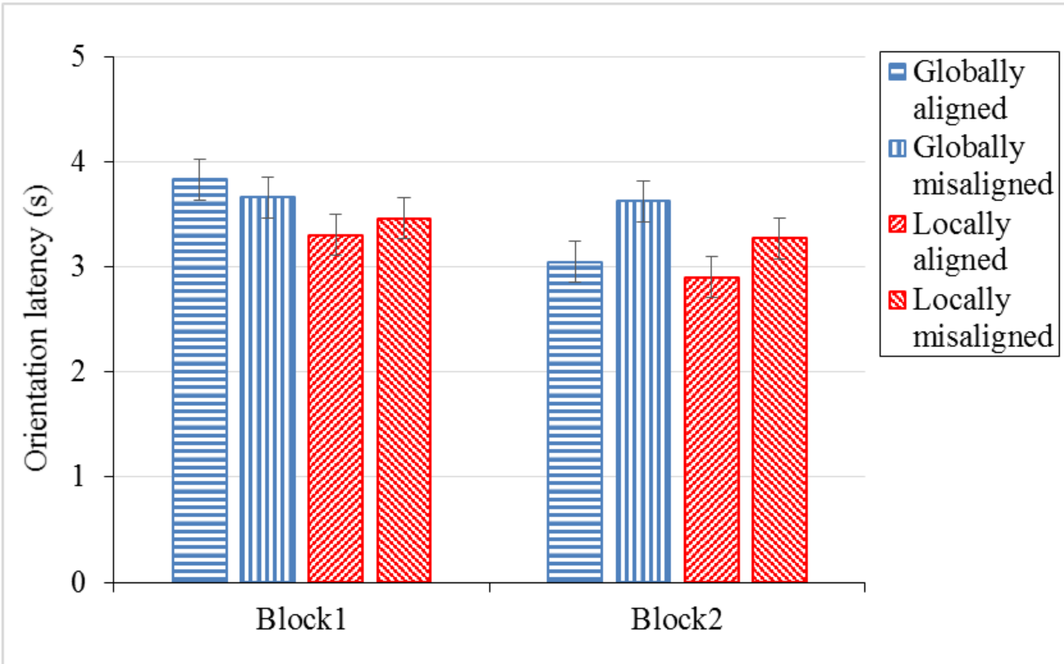

Figure S3. The mean orientation latency for four types of trials in two blocks of the priming task in the replication experiment. Error bars represent ±1 SE removing the variance from individual differences.

1.5 Discussion

The results showed both global and local priming effects using the trials adapted from the remote perspective-taking trials in the current study, which replicated the findings in Lei et al. (2019). More importantly, these results indicated that the experimental setup and the learning procedure used in Experiments 1, 2, and 3 in the current study could develop global and local representations.

2. Results from orientation latency in Experiments 1, 2, and 3

There were no significant results from orientation latency in the remote perspective-taking task for either the global or local sensorimotor alignment effect in any experiment (Figure S4).

In Experiment 1, there was no significant difference between the globally aligned and globally misaligned conditions, $t(47) = 0.24$, $p = .815$, Cohen's $d = 0.05$, $BF_{01} = 8.62$. This result shows no evidence for a global sensorimotor alignment effect. There was no significant difference between the locally aligned and locally misaligned conditions, $t(47) = 1.79$, $p = .080$, Cohen's $d = 0.37$, $BF_{01} = 1.93$, indicating no clear evidence for a local sensorimotor alignment effect.

In Experiment 2, there was no significant difference between the globally aligned and globally misaligned conditions, $t(47) = 0.48$, $p = .634$, Cohen's $d = 0.10$, $BF_{01} = 7.91$, indicating no global sensorimotor alignment effect. There was no significant difference between the locally aligned and locally misaligned conditions, $t(47) = 0.21$, $p = .831$, Cohen's $d = 0.04$, $BF_{01} = 8.66$, indicating no local sensorimotor alignment effect.

In Experiment 3, there was no significant difference between the globally aligned and globally misaligned conditions, $t(31) = 0.46$, $p = .652$, Cohen's $d = 0.11$, $BF_{01} = 6.60$, indicating no global sensorimotor alignment effect. There was no significant difference between the locally aligned and locally misaligned conditions, $t(31) = 1.85$, $p = .074$, Cohen's $d = 0.46$, $BF_{01} = 1.50$, indicating no local sensorimotor alignment effect.

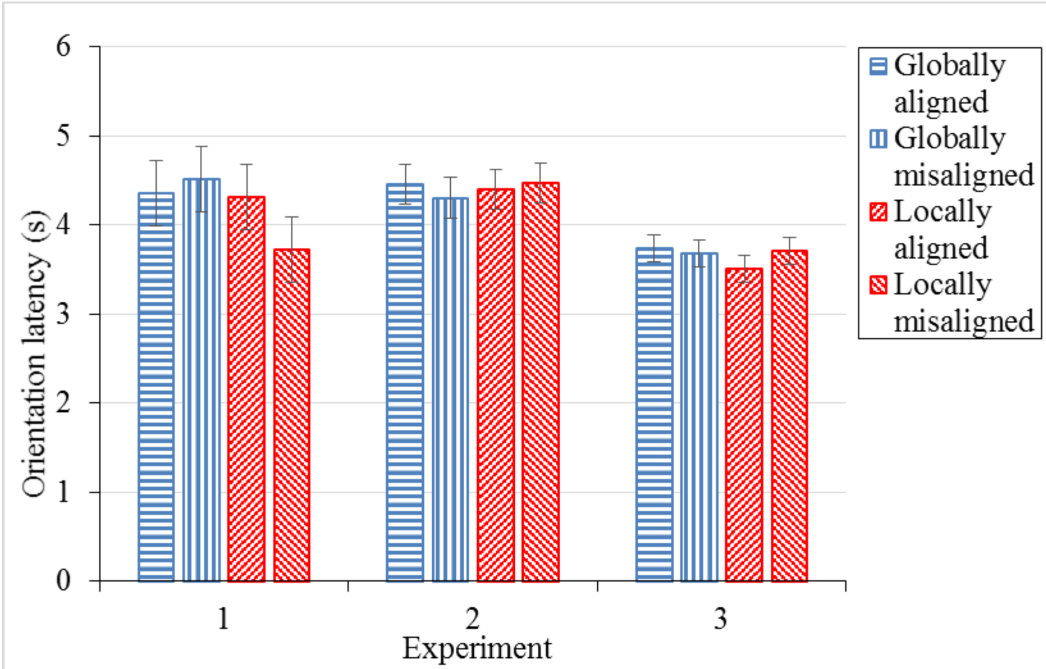

Figure S4. The mean orientation latency for four types of trials in the remote perspective-taking task in all experiments. Error bars represent ±1 SE removing the variance from individual differences.

3. Relative heading judgment task in Experiment 2

3.1 Data analysis

In the relative heading judgment task, the participants' heading responses were recorded. We calculated the signed response errors by subtracting the heading responses from the direction of the probed views. The signed response errors in the within-room trials reflected the accuracy of the local representations in the two rooms and also whether the participants understood and conducted the task correctly. The signed response errors in the across-room trials reflected the accuracy of the global representations. The participants' global representations might have been affected by the local representations so that the locally consistent views appeared to be globally closer. For example, Views 1 and 9 were globally 90º different. However, the participants might have represented the angular difference between them to be smaller than 90º. Therefore, when the actual views were in Room 1 (Figure 1) and the probed views were in Room 2, the signed response errors due to this bias would be counterclockwise (or negative). When the actual views were in Room 2 and the probed views were in Room 1, the signed response errors due to this bias would be clockwise (or positive).

To examine this local bias for the across-room trials, we reversed the sign of the signed response error when the actual view was in Room 2, so that the response bias from the local representations was also counterclockwise, which would be the same as when the actual view was in Room 1. Hence, if the signed response error was close to 0º, it would indicate accurate global representations, whereas if the signed response error was counterclockwise, it would indicate the influence of the local representations on the global representations. The signed response errors in the within-room trials were not changed since these trials tested the local representations and thus there was no such issue of local bias. Then we calculated the circular means of the signed response errors for the across-room and within-room trials. The mean response latencies were also calculated for these two types of trials.

3.2 Results

3.2.1 Signed response error

Figure S5 shows the signed response errors in the across-room and within-room trials. The circular mean in the within-room trials was 2.66º and the 95% confidence interval covered 0º, indicating that the participants had accurate local representations. The circular mean in the across-room trials was 304.76º and the 95% confidence interval did not cover 0º. This indicates that the global representations were not as accurate as the local representations and were influenced by the local representations.

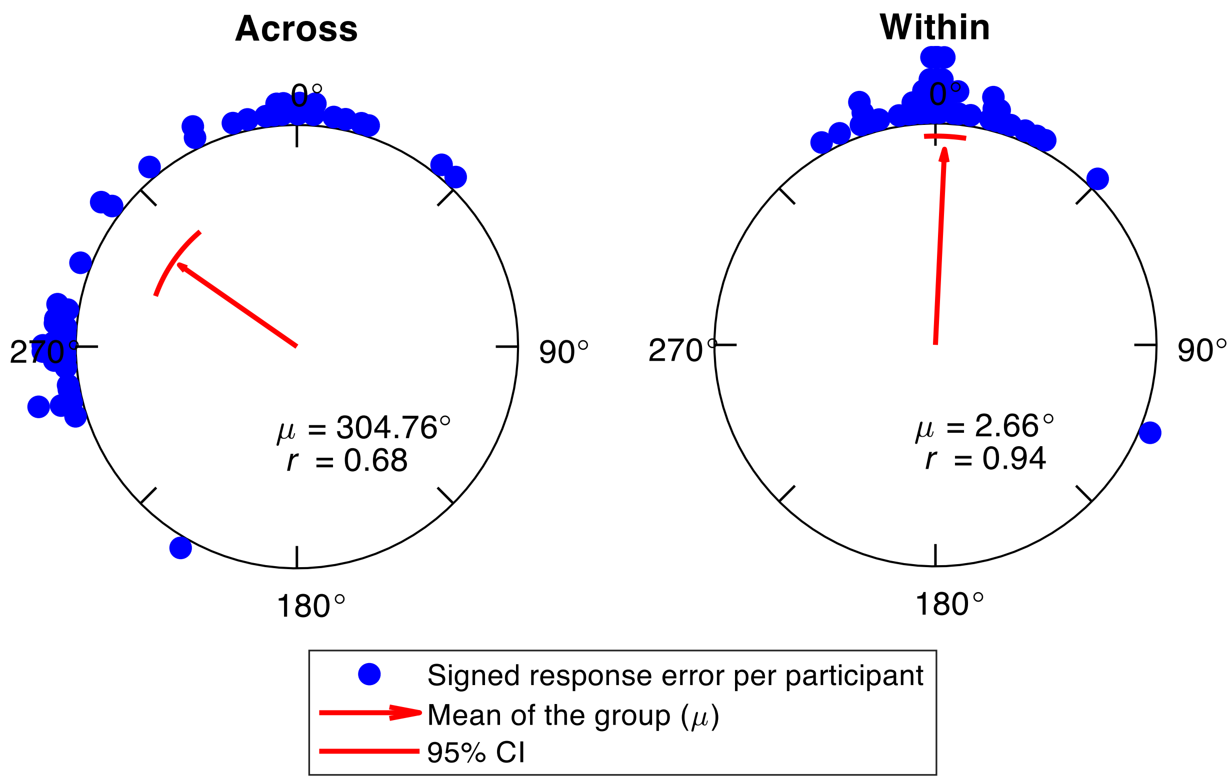

Figure S5. The signed response errors in the across-room and within-room trials in the relative heading judgment task in Experiment 2. Each dot represents the error for one participant. The red arrow indicates the circular direction (μ) and the length (r) of the mean vector of the signed response errors across the participants. The arc above the arrow indicates the 95% confidence interval of the circular mean direction.

3.2.2 Response latency

The response latency in the within-room trials (M= 11.83s, SD= 7.28s) was significantly faster than that in the across-room trials (M= 15.03s, SD= 7.64s), $t(47) = 3.29$, $p = .002$, Cohen's $d = 0.67$. This might be because the local representations were more coherent and thus were accessed more easily and quickly, but the global representations were more piecewise and thus required more time to calculate the heading response across rooms.

3.2.3 Correlation with the global sensorimotor alignment effect

We examined whether the global performances in the relative heading judgment task and the remote perspective-taking task were correlated. We calculated the correlation between the absolute response errors in the across-room trials of the relative heading judgment and the absolute pointing errors of the global sensorimotor alignment effect (i.e., differences between the globally aligned and misaligned trials) in the remote perspective-taking task. There was no significant correlation, $r(46)=-0.091$, $p=0.539$.

4. Relative heading judgment task in Experiment 3

4.1 Results

The data analysis was the same as in Experiment 2.

4.1.1 Signed response error

Figure S6 shows the signed response errors in the across-room and within-room trials. The circular mean in the within-room trials was 359.08º and the 95% confidence interval covered 0º, indicating that the participants had accurate local representations. The circular mean in the across-room trials was 292.16º and the 95% confidence interval did not cover 0º. This indicates that the global representations were not as accurate as the local representations and were influenced by the local representations.

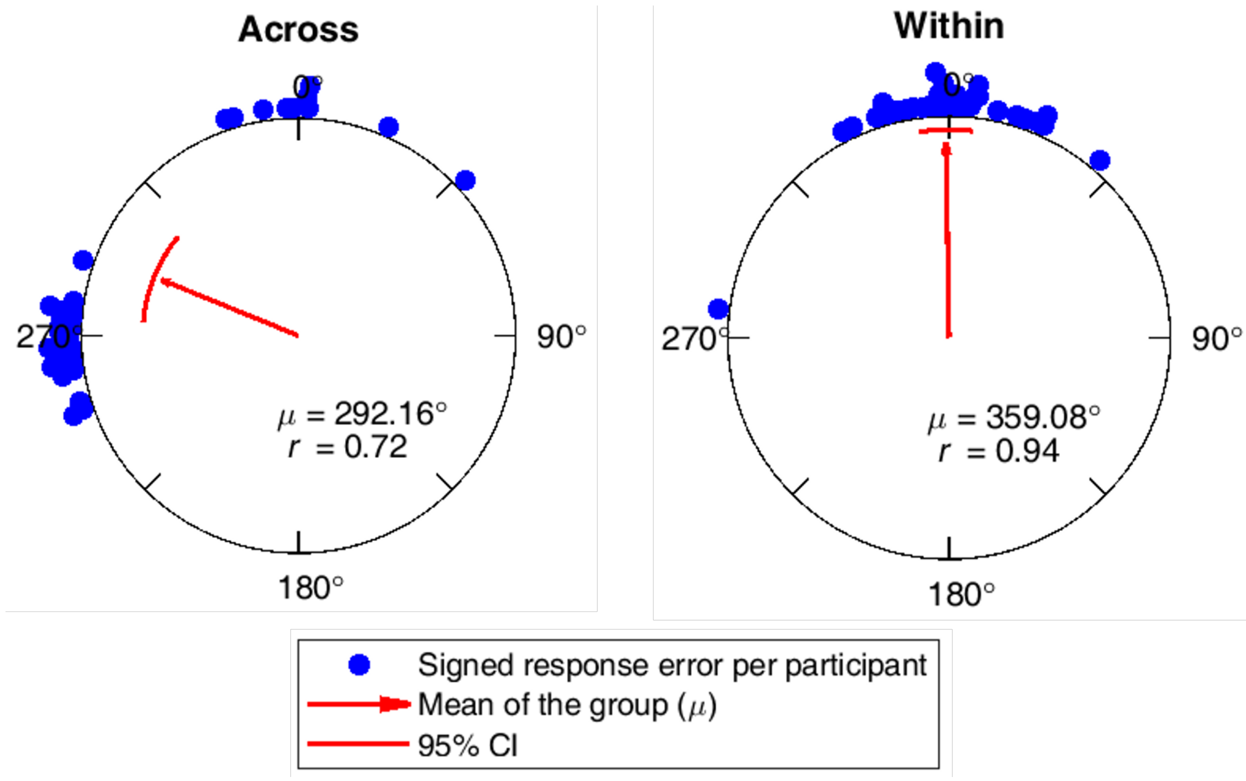

Figure S6. The signed response errors in the across-room and within-room trials in the relative heading judgment task in Experiment 3. Each dot represents the error for one participant. The red arrow indicates the circular direction (μ) and the length (r) of the mean vector of the signed response errors across the participants. The arc above the arrow indicates the 95% confidence interval of the circular mean direction.

4.1.2 Response latency

The response latency in the within-room trials (M= 12.42s, SD= 7.18s) was not different from that in the across-room trials (M= 10.93s, SD= 6.75s), $t(31) = 1.03$, $p = .313$, Cohen's d = 0.26, BF_{01}=4.41.

4.1.3 Correlation with the global sensorimotor alignment effect

We calculated the correlation between the absolute response errors in the across-room trials of the relative heading judgment and the absolute pointing errors of the global sensorimotor alignment effect (i.e., differences between the globally aligned and misaligned trials) in the remote perspective-taking task. There was no significant correlation, $r(30)=-0.105$, $p=0.566$.